Activity Report 2012

Project-Team PARIETAL

Modelling brain structure, function and variability based on high-field MRI data.

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Medicine and Neuro-sciences

Table of contents

Project-Team PARIETAL

Keywords: Medical Images, Image Processing, Biological Images, Brain Computer Interface, Machine Learning

PARIETAL is an Inria Research Team within the Neurospin platform of CEA Institute. It is located in Saclay, in Neurospin building. Parietal aims at addressing several issues raised by the analysis of high-resolution neuroimaging data, and gives a free access to its tools for brain image analysis and machine learning.

Creation of the Project-Team: July 01, 2009 .

1. Members

Research Scientists

Bertrand Thirion [Team Leader, Ingénieur spécialiste, Inria, HdR]
Jean-Baptiste Poline [CEA, Neurospin, HdR]
Pierre Fillard [Junior Researcher]
Gaël Varoquaux [Junior Researcher]
Stéphanie Allassonnière [Associate Professor, collaborateur extérieur]
Alexandre Gramfort [Associate Professor, collaborateur extérieur]
Dimitris Samaras [Associate Professor, collaborateur extérieur]

Engineers

Sergio Medina [ADT]
Jaques Grobler [ADT]
Elvis Dohmatob [ADT]

PhD Students

Virgile Fritsch [Digiteo grant]
Viviana Siless [MESR grant]
Fabian Pedregosa [ANR grant]
Michael Eickenberg [MESR grant]
Yannick Schwartz [ANR grant]
Hao Xu [collaborateur extérieur]
Alexandre Abraham [Digiteo grant]

Post-Doctoral Fellows

Benoit da Mota [Digiteo grant]
Bernard Ng [DRI grant]

Administrative Assistant

Régine Bricquet

2. Overall Objectives

2.1. Highlights of the Year

Fabian Pedregosa, PhD candidate at the Parietal team won the best poster award at the EuroScipy 2012 conference. The poster, Memory Profiler: monitor memory usage of Python code describes the Python package memory_profiler, a tool to monitor memory usage from within the Python language. Among other features, the package is able to perform line-by-line analysis of the memory usage program and to insert breakpoints on excessive memory consumption.

3. Scientific Foundations

3.1. Human neuroimaging data and its use

Human neuroimaging consists in acquiring non-invasively image data from normal and diseased human populations. Magnetic Resonance Imaging (MRI) can be used to acquire information on brain structure and function at high spatial resolution.

- T1-weighted MRI is used to obtain a segmentation of the brain into different different tissues, such as gray matter, white matter, deep nuclei, cerebro-spinal fluid, at the millimeter or sub-millimeter resolution. This can then be used to derive geometric and anatomical information on the brain, e.g. cortical thickness.
- Diffusion-weighted MRI measures the local diffusion of water molecules in the brain at the resolution of 2mm, in a set of directions (30 to 60 typically). Local anisotropy, observed in white matter, yields a geometric model of fiber tracts along which water diffusion occurs, and thus provides essential information of the connectivity structure of the brain.
- Functional MRI measures the blood-oxygen-level-dependent (BOLD) contrast that reflects neural activity in the brain, at a spatial resolution of 2 to 3mm, and a temporal resolution of 2-3s. This yields a spatially resolved image of brain functional networks that can be modulated either by specific cognitive tasks or appear as networks of correlated activity.
- Electro- and Magneto-encephalography (MEEG) are two additional modalities that complement functional MRI, as they directly measure the electric and magnetic signals elicited by neural activity, at the millisecond scale. These modalities rely on surface measurements and do not localize brain activity very accurately in the spatial domain.

3.2. High-field MRI

High field MRI as performed at Neurospin (7T on humans, 11.7T in 2013, 17.6T on rats) brings an improvement over traditional MRI acquisitions at 1.5T or 3T, related to to a higher signal-to-noise ratio in the data. Depending on the data and applicative context, this gain in SNR can be traded against spatial resolution improvements, thus helping in getting more detailed views of brain structure and function. This comes at the risk of higher susceptibility distortions of the MRI scans and signal inhomogeneities, that need to be corrected for. Improvements at the acquisition level may come from the use of new coils (such as the new 32 channels coil on the 7T at Neurospin).

3.3. Technical challenges for the analysis of neuroimaging data

The first limitation of Neuroimaging-based brain analysis is the limited Signal-to-Noise Ratio of the data. A particularly striking case if functional MRI, where only a fraction of the data is actually understood, and from which it is impossible to observe by eye the effect of neural activation on the raw data. Moreover, far from traditional i.i.d. Gaussian models, the noise in MRI typically exhibits correlations and long-distance correlation properties (e.g. motion-related signal) and has potentially large amplitude, which can make it hard to distinguish from true signal on a purely statistical basis. A related difficulty is the *lack of salient structure* in the data: it is hard to infer meaningful patterns (either through segmentation or factorization procedures) based on the data only. A typical case is the inference of brain networks from resting-state functional connectivity data.

Regarding statistical methodology, neuroimaging problems also suffer from the relative paucity of the data, i.e. the relatively small number of images available to learn brain features or models, e.g. with respect to the size of the images or the number of potential structures of interest. This leads to several kinds of difficulties, known either as multiple comparison problems or curse of dimensionality. One possibility to overcome this challenge is to increase the amount of data by using images from multiple acquisition centers, at the risk of introducing scanner-related variability, thus challenging the homogeneity of the data. This becomes an important concern with the advent of cross-modal neuroimaging-genetics studies.

4. Application Domains

4.1. Inverse problems in Neuroimaging

Many problems in neuroimaging can be framed as forward and inverse problems. For instance, the neuroimaging *inverse problem* consists in predicting individual information (behavior, phenotype) from neuroimaging data, while an important the *forward problem* consists in fitting neuroimaging data with high-dimensional (e.g. genetic) variables. Solving these problems entails the definition of two terms: a loss that quantifies the goodness of fit of the solution (does the model explain the data reasonably well ?), and a regularization schemes that represents a prior on the expected solution of the problem. In particular some priors enforce some properties of the solutions, such as sparsity, smoothness or being piecewise constant.
Let us detail the model used in the inverse problem: Let X be a neuroimaging dataset as an (n_{subj}, n_{voxels}) matrix, where n_{subj} and n_{voxels} are the number of subjects under study, and the image size respectively, Y an array of values that represent characteristics of interest in the observed population, written as (n_{subj}, n_f) matrix, where n_f is the number of characteristics that are tested, and β an array of shape (n_{voxels}, n_f) that represents a set of pattern-specific maps. In the first place, we may consider the columns $Y_1, .., Y_{n_f}$ of Y independently, yielding n_f problems to be solved in parallel:

$$Y_i = X\beta_i + \epsilon_i, \forall i \in \{1, .., n_f\},$$

where the vector contains β_i is the i^{th} line of β. As the problem is clearly ill-posed, it is naturally handled in a regularized regression framework:

$$\widehat{\beta}_i = \text{argmin}_{\beta_i} \|Y_i - X\beta_i\|^2 + \Psi(\beta_i), \tag{1}$$

where Ψ is an adequate penalization used to regularize the solution:

$$\Psi(\beta; \lambda_1, \lambda_2, \eta_1, \eta_2) = \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2^2 + \eta_1 \|\nabla\beta\|_1 + \eta_2 \|\nabla\beta\|_2^2 \tag{2}$$

with $\lambda_1,\ \lambda_2,\ \eta_1,\ \eta_2 \geq 0$. In general, only one or two of these constraints will have a non-zero weighting:

- When $\lambda_1 > 0$ only (LASSO), and to some extent, when $\lambda_1, \lambda_2 > 0$ only (elastic net), the optimal solution β is (possibly very) sparse, but may not exhibit a proper image structure; it does not fit well with the intuitive concept of a brain map.
- Total Variation regularization (see Fig. 1) is obtained for ($\eta_1 > 0$ only), and typically yields a piecewise constant solution.
- Smooth lasso is obtained with ($\eta_2 > 0$ and $\lambda_1 > 0$ only), and yields smooth, compactly supported spatial basis functions.

The performance of the predictive model can simply be evaluated as the amount of variance in Y_i fitted by the model, for each $i \in \{1, .., n_f\}$. This can be computed through cross-validation, by *learning* $\widehat{\beta}_i$ on some part of the dataset, and then estimating $(Y_i - X\widehat{\beta}_i)$ using the remainder of the dataset.

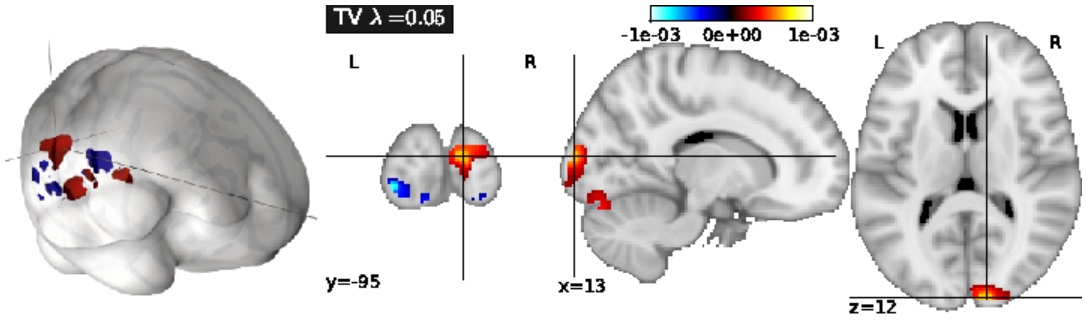


Figure 1. Example of the regularization of a brain map with total variation in an inverse problem. The problems here consists in predicting the spatial scale of an object presented as a stimulus, given functional neuroimaging data acquired during the observation of an image. Learning and test are performed across individuals. Unlike other approaches, Total Variation regularization yields a sparse and well-localized solution that enjoys particularly high accuracy.

This framework is easily extended by considering

- *Grouped penalization*, where the penalization explicitly includes a prior clustering of the features, i.e. voxel-related signals, into given groups. This is particularly important to include external anatomical priors on the relevant solution.
- *Combined penalizations*, i.e. a mixture of simple and group-wise penalizations, that allow some variability to fit the data in different populations of subjects, while keeping some common constraints.
- *Logistic regression*, where a sigmoid non-linearity is applied to the linear model so that it yields a probability of classification in a binary classification problem.
- *Robustness to between-subject variability* is an important question, as it makes little sense that a learned model depends dramatically on the particular observations used for learning. This is an important issue, as this kind of robustness is somewhat opposite to sparsity requirements.
- *Multi-task learning*: if several target variables are thought to be related, it might be useful to constrain the estimated parameter vector β to have a shared support across all these variables.
  For instance, when one of the variables Y_i is not well fitted by the model, the estimation of other variables $Y_j, j \neq i$ may provide constraints on the support of β_i and thus, improve the prediction of Y_i. Yet this does not impose constraints on the non-zero parameters of the parameters β_i.

$$Y = X\beta + \epsilon, \tag{3}$$

then

$$\widehat{\beta} = \text{argmin}_{\beta=(\beta_i),i=1..n_f} \sum_{i=1}^{n_f} \|Y_i - X\beta_i\|^2 + \lambda \sum_{j=1}^{n_{voxels}} \sqrt{\sum_{i=1}^{n_f} \beta_{i,j}^2} \tag{4}$$

4.2. Multivariate decompositions

Multivariate decompositions are an important tool to model complex data such as brain activation images: for instance, one might be interested in extracting an atlas of brain regions from a given dataset, such as regions depicting similar activities during a protocol, across multiple protocols, or even in the absence of protocol (during resting-state). These data can often be factorized into spatial-temporal components, and thus can be estimated through *regularized Principal Components Analysis* (PCA) algorithms, which share some common steps with regularized regression.
Let X be a neuroimaging dataset written as an (n_{subj}, n_{voxels}) matrix, after proper centering; the model reads

$$X = AD + \epsilon, \tag{5}$$

where D represents a set of n_{comp} spatial maps, hence a matrix of shape (n_{comp}, n_{voxels}), and A the associated subject-wise loadings. While traditional PCA and independent components analysis are limited to reconstruct components D within the space spanned by the column of X, it seems desirable to add some constraints on the rows of D, that represent spatial maps, such as sparsity, and/or smoothness, as it makes the interpretation of these maps clearer in the context of neuroimaging.
This yields the following estimation problem:

$$\min_{D,A} \|X - AD\|^2 + \Psi(D) \text{ s.t. } \|A_i\| = 1 \; \forall i \in \{1..n_f\}, \tag{6}$$

where $(A_i),\ i \in \{1..n_f\}$ represents the columns of A. Ψ can be chosen such as in Eq. (2) in order to enforce smoothness and/or sparsity constraints.
The problem is not jointly convex in all the variables but each penalization given in Eq (2) yields a convex problem on D for A fixed, and conversely. This readily suggests an alternate optimization scheme, where D and A are estimated in turn, until convergence to a local optimum of the criterion. As in PCA, the extracted components can be ranked according to the amount of fitted variance. Importantly, also, estimated PCA models can be interpreted as a probabilistic model of the data, assuming a high-dimensional Gaussian distribution (probabilistic PCA).

4.3. Covariance estimation

Another important estimation problem stems from the general issue of learning the relationship between sets of variables, in particular their covariance. Covariance learning is essential to model the dependence of these variables when they are used in a multivariate model, for instance to assess whether an observation is aberrant or not or in classification problems. Covariance learning is necessary to model latent interactions in high-dimensional observation spaces, e.g. when considering multiple contrasts or functional connectivity data.
The difficulties are two-fold: on the one hand, there is a shortage of data to learn a good covariance model from an individual subject, and on the other hand, subject-to-subject variability poses a serious challenge to the use of multi-subject data. While the covariance structure may vary from population to population, or depending on the input data (activation versus spontaneous activity), assuming some shared structure across problems, such as their sparsity pattern, is important in order to obtain correct estimates from noisy data. Some of the most important models are:

- **Sparse Gaussian graphical models**, as they express meaningful conditional independence relationships between regions, and do improve conditioning/avoid overfit.
- **Decomposable models**, as they enjoy good computational properties and enable intuitive interpretations of the network structure. Whether they can faithfully or not represent brain networks is an important question that needs to be addressed.
- **PCA-based regularization of covariance** which is powerful when modes of variation are more important than conditional independence relationships.

Adequate model selection procedures are necessary to achieve the right level of sparsity or regularization in covariance estimation; the natural evaluation metric here is the out-of-samples likelihood of the associated Gaussian model. Another essential remaining issue is to develop an adequate statistical framework to test differences between covariance models in different populations. To do so, we will consider different means of parametrizing covariance distributions and their impact on the network. Our current work on post-stroke patients (see e.g. Fig. 2) suggests indeed that modeling may prove essential to perform sensitive inference.

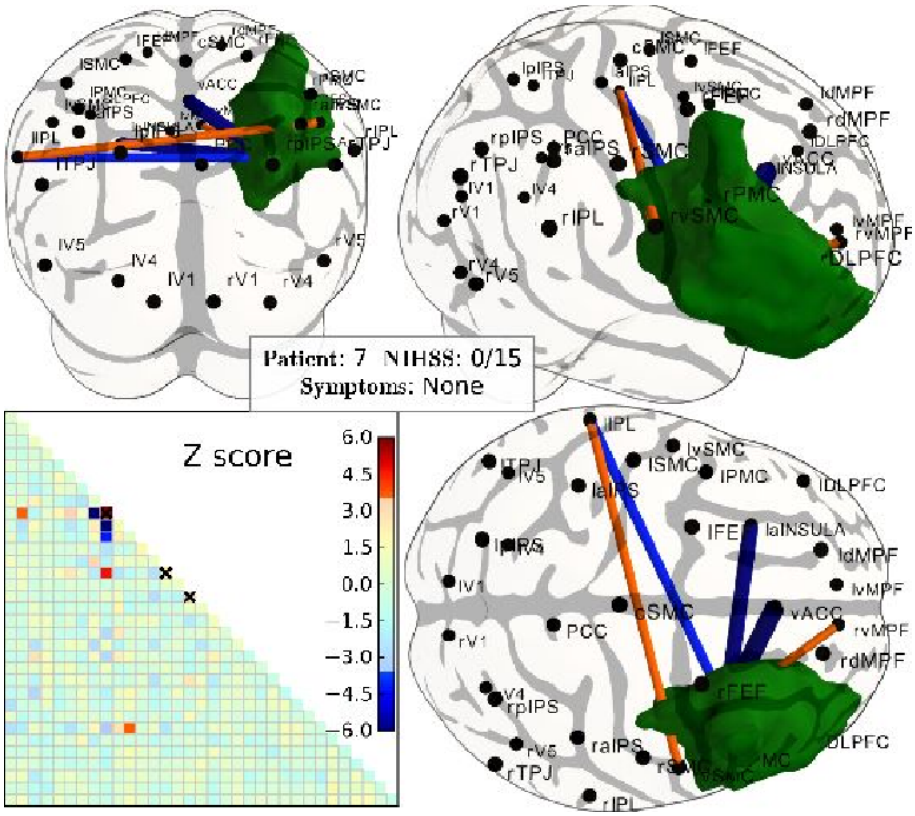


Figure 2. Example of functional connectivity analysis: The correlation matrix describing brain functional connectivity in a post-stroke patient (lesion outlined in green) is compared to a group of control subjects. Some edges of the graphical model show a significant difference, but the statistical detection of the difference requires a sophisticated statistical framework for the comparison of graphical models.

5. Software

5.1. Mayavi

Participant: Gaël Varoquaux [Correspondant].

Mayavi is the most used scientific 3D visualization Python software (http://mayavi.sourceforge.net/). It has been developed by Prabhu Ramachandran (IIT Bombay) and Gaël Varoquaux (PARIETAL, Inria Saclay). Mayavi can be used as a visualization tool, through interactive command line or as a library. It is distributed under Linux through Ubuntu, Debian, Fedora and Mandriva, as well as in PythonXY and EPD Python scientific distributions. Mayavi is used by several software platforms, such as PDE solvers (fipy, sfepy), molecule visualization tools (http://pyrx.scripps.edu) and brain connectivity analysis tools (connectomeViewer).

See also the web page http://mayavi.sourceforge.net/ and the following paper http://hal.inria.fr/inria-00528985/en.

- Version: 3.4.0

5.2. Nipy

Participants: Bertrand Thirion [correspondant], Virgile Fritsch, Elvis Dohmatob, Gaël Varoquaux.

Nipy is an open-source Python library for neuroimaging data analysis, developed mainly at Berkeley, Stanford, MIT and Neurospin. It is open to any contributors and aims at developing code and tools sharing. Some parts of the library are completely developed by Parietal and LNAO (CEA, DSV, Neurospin). It is devoted to algorithmic solutions for various issues in neuroimaging data analysis. All the nipy project is freely available, under BSD license. It is available in NeuroDebian.

See also the web page http://nipy.org.

- Version: 0.3

5.3. MedInria

Participants: Pierre Fillard [correspondant], Sergio Medina, Viviana Siless.

MedInria is a free collection of softwares developed within the ASCLEPIOS, ATHENA and VISAGES research projects. It aims at providing to clinicians state-of-the-art algorithms dedicated to medical image processing and visualization. Efforts have been made to simplify the user interface, while keeping high-level algorithms. MedInria is available for Microsoft windows XP/Vista, Linux Fedora Core, MacOSX, and is fully multi-threaded.
See also the web page http://med.inria.fr/.

- Version: 2.0

5.4. Scikit learn

Participants: Bertrand Thirion, Gaël Varoquaux [correspondant], Jaques Grobler, Alexandre Gramfort, Fabian Pedregosa, Virgile Fritsch.

Scikit-learn is an open-source machine learning toolkit written in Python/C that provides generic tools to learn information for the classification of various kinds of data, such as images or texts. It is tightly associated to the scientific Python software suite (numpy/scipy) for which it aims at providing a complementary toolkit for machine learning (classification, clustering, dimension reduction, regression). There is an important focus on code quality (API consistency, code readability, tests, documentation and examples), and on efficiency, as the scikit-learn compares favorably to state-of-the-art modules developed in R in terms of computation time or memory requirements. Scikit-learn is currently developed by more than 60 contributors, but the core developer team has been with the Parietal Inria team at Saclay-Île-de-France since January 2010. The scikit-learn has recently become the reference machine learning library in Python.

- Version: 0.12
- Programming language: Python, C/Cython

6. New Results

6.1. Randomized cluster-based predictive model

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Alexandre Gramfort.

Functional neuroimaging can measure the brain's response to an external stimulus. It is used to perform brain mapping: identifying from these observations the brain regions involved. This problem can be cast into a linear supervised learning task where the neuroimaging data are used as predictors for the stimulus. Brain mapping is then seen as a support recovery problem. On functional MRI (fMRI) data, this problem is particularly challenging as i) the number of samples is small due to limited acquisition time and ii) the variables are strongly correlated. We propose to overcome these difficulties using sparse regression models over new variables obtained by clustering of the original variables. The use of randomization techniques, e.g. bootstrap samples, and clustering of the variables improves the recovery properties of sparse methods. We demonstrate the benefit of our approach on an extensive simulation study as well as two fMRI datasets.

More details can be found in [32].

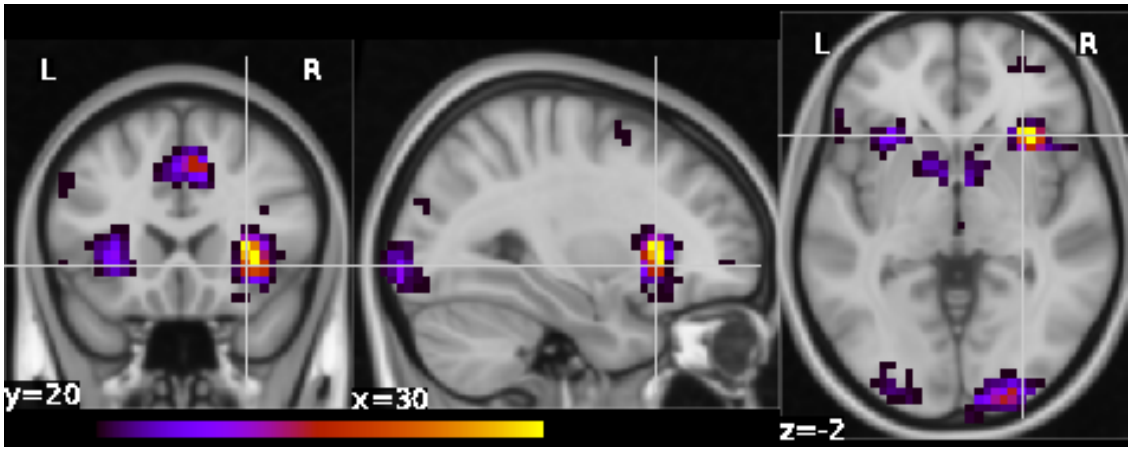


Figure 3. The randomized cluster-based predictive model can be used to predict the behavior of the subject, such as the gain in a gambling task [33]. More importantly, the support of the resulting classifier is indeed sparse and provides a reliable definition of the truly involved regions.

6.2. Random Projections for Outlier Detection

Participants: Gaël Varoquaux, Bertrand Thirion, Jean-Baptiste Poline, Virgile Fritsch [Correspondant].

Medical imaging datasets often contain deviant observations, the so-called outliers, due to acquisition or preprocessing artifacts or resulting from large intrinsic inter-subject variability. These can undermine the statistical procedures used in group studies as the latter assume that the cohorts are composed of homogeneous samples with anatomical or functional features clustered around a central mode. The effects of outlying subjects can be mitigated by detecting and removing them with explicit statistical control. With the emergence of large medical imaging databases, exhaustive data screening is no longer possible, and automated outlier detection methods are currently gaining interest. The datasets used in medical imaging are often high-dimensional and strongly correlated. The outlier detection procedure should therefore rely on high-dimensional statistical multivariate models. However, state-of-the-art procedures are not well-suited for such high-dimensional

settings. In this work, we introduce regularization in the Minimum Covariance Determinant framework and investigate different regularization schemes. We carry out extensive simulations to provide backing for practical choices in absence of ground truth knowledge. We demonstrate on functional neuroimaging datasets that outlier detection can be performed with small sample sizes and improves group studies.

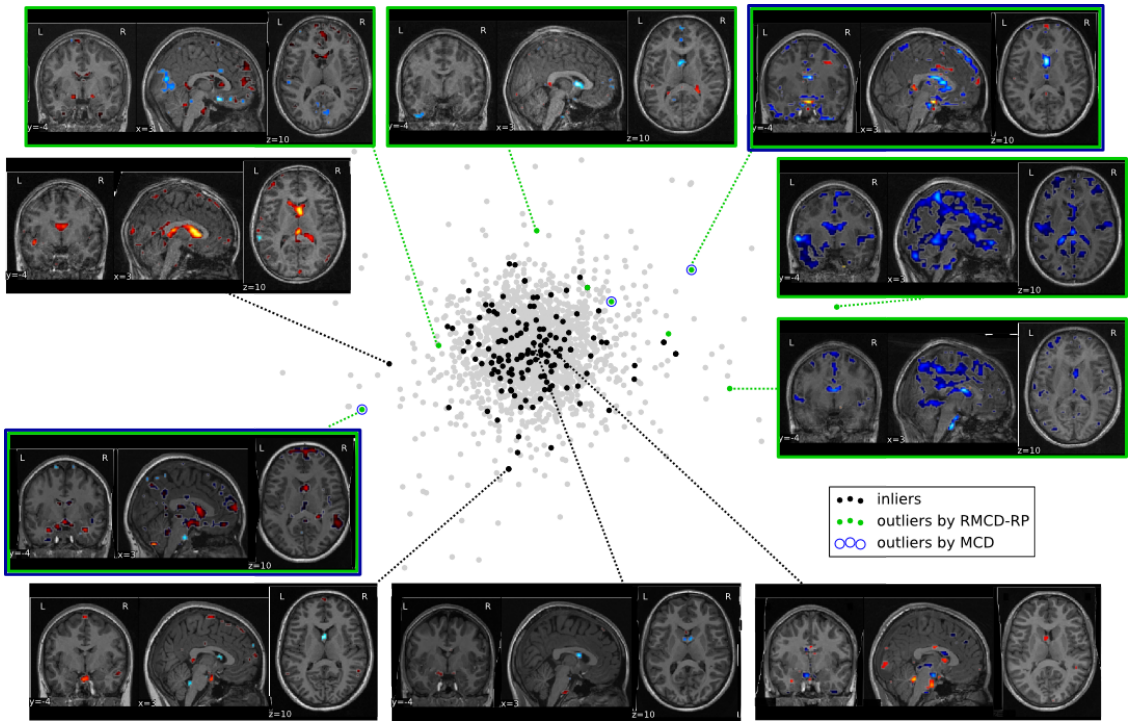


Figure 4. A large set of images can be mined for structures using the regularized MCD framework, which reveals both standard and unusual patterns in these images.

More details can be found in [11].

6.3. Registration of brain images based on Currents

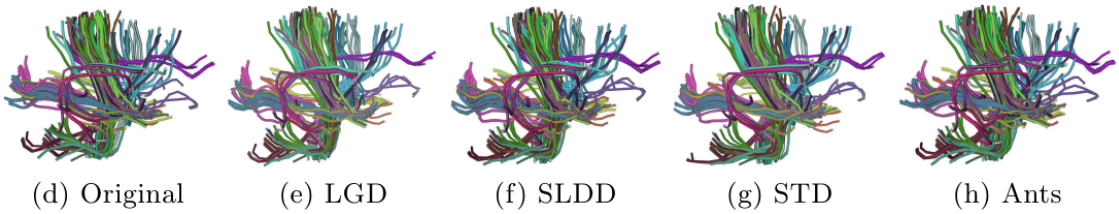
Participants: Pierre Fillard, Bertrand Thirion, Viviana Siless [correspondant].

We present an extension of the diffeomorphic Geometric Demons algorithm which combines the iconic registration with geometric constraints. Our algorithm works in the log-domain space, so that one can efficiently compute the deformation field of the geometry. We represent the shape of objects of interest in the space of currents which is sensitive to both location and geometric structure of objects. Currents provide a distance between geometric structures that can be defined without specifying explicit point-to-point correspondences. We demonstrate this framework by registering simultaneously T1 images and 65 fiber bundles consistently extracted in 12 subjects and compare it against non-linear T1, tensor, and multi-modal T1+ Fractional Anisotropy (FA) registration algorithms. Results show the superiority of the Log-domain Geometric Demons over their purely iconic counterparts.

More details can be found in [31].

6.4. Structured Sparsity for brain mapping

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Alexandre Gramfort.



(d) Original (e) LGD (f) SLDD (g) STD (h) Ants

Figure 5. Comparison of the fiber registration through various algorithms. We display a moving and a reference fiber for 29 selected bundles. The proposed approach (SLDD) outperforms state-of-the art alternatives that do not take into account the fiber geometry explicitly.

Reverse inference, or *brain reading*, is a recent paradigm for analyzing functional magnetic resonance imaging (fMRI) data, based on pattern recognition and statistical learning. This approach aims at decoding brain activity by predicting some cognitive variables related to brain activation maps. Reverse inference takes into account the multivariate information between voxels and is currently the only way to assess how precisely some cognitive information is encoded by the activity of neural populations within the whole brain. However, it relies on a prediction function that is plagued by the curse of dimensionality, since there are far more features than samples, i.e., more voxels than fMRI volumes. To address this problem, different methods have been proposed, such as, among others, univariate feature selection, feature agglomeration and regularization techniques. In this work, we consider a sparse hierarchical structured regularization. Specifically, the penalization we use is constructed from a tree that is obtained by spatially-constrained agglomerative clustering. This approach encodes the spatial structure of the data at different scales into the regularization, which makes the overall prediction procedure more robust to inter-subject variability. The regularization used induces the selection of spatially coherent predictive brain regions simultaneously at different scales. We test our algorithm on real data acquired to study the mental representation of objects, and we show that the proposed algorithm not only delineates meaningful brain regions but yields as well better prediction accuracy than reference methods.

More details can be found in [15].

6.5. A Novel Sparse Graphical Approach for Multimodal Brain Connectivity Inference

Participants: Bertrand Thirion, Jean-Baptiste Poline, Gaël Varoquaux [Correspondant], Bernard Ng.

Despite the clear potential benefits of combining fMRI and diffusion MRI in learning the neural pathways that underlie brain functions, little methodological progress has been made in this direction. In this work, we propose a novel multimodal integration approach based on sparse Gaussian graphical model for estimating brain connectivity. Casting functional connectivity estimation as a sparse inverse covariance learning problem, we adapt the level of sparse penalization on each connection based on its anatomical capacity for functional interactions. Functional connections with little anatomical support are thus more heavily penalized. For validation, we showed on real data collected from a cohort of 60 subjects that additionally modeling anatomical capacity significantly increases subject consistency in the detected connection patterns. Moreover, we demonstrated that incorporating a connectivity prior learned with our multimodal connectivity estimation approach improves activation detection.

More details can be found in [26].

6.6. Transfer learning for met-analyses of functional neuroimaging datasets

Participants: Bertrand Thirion, Jean-Baptiste Poline, Gaël Varoquaux, Yannick Schwartz [Correspondant].

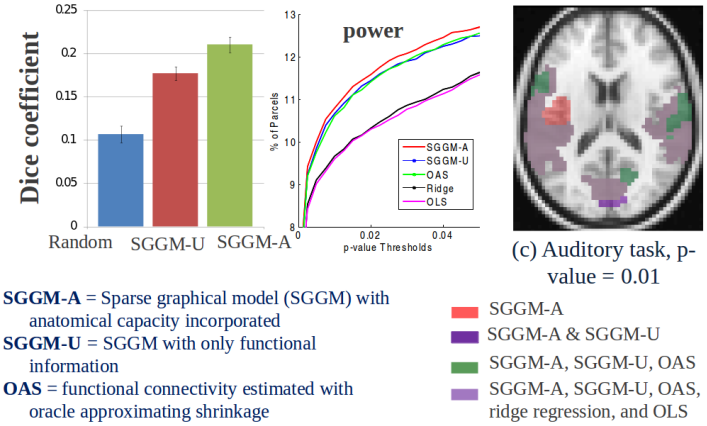


Figure 6. The information conveyed by anatomical connectivity improves the estimation of functional connectivity, as it makes it more reproducible. It also enhances the power of fMRI activation detection studies when used as a prior on these activation maps.

Typical cohorts in brain imaging studies are not large enough for systematic testing of all the information contained in the images. To build testable working hypotheses, investigators thus rely on analysis of previous work, sometimes formalized in a so-called meta-analysis. In brain imaging, this approach underlies the specification of regions of interest (ROIs) that are usually selected on the basis of the coordinates of previously detected effects. In this work, we propose to use a database of images, rather than coordinates, and frame the problem as transfer learning: learning a discriminant model on a reference task to apply it to a different but related new task. To facilitate statistical analysis of small cohorts, we use a sparse discriminant model that selects predictive voxels on the reference task and thus provides a principled procedure to define ROIs. The benefits of our approach are twofold. First it uses the reference database for prediction, i.e. to provide potential biomarkers in a clinical setting. Second it increases statistical power on the new task. We demonstrate on a set of 18 pairs of functional MRI experimental conditions that our approach gives good prediction. In addition, on a specific transfer situation involving different scanners at different locations, we show that voxel selection based on transfer learning leads to higher detection power on small cohorts.

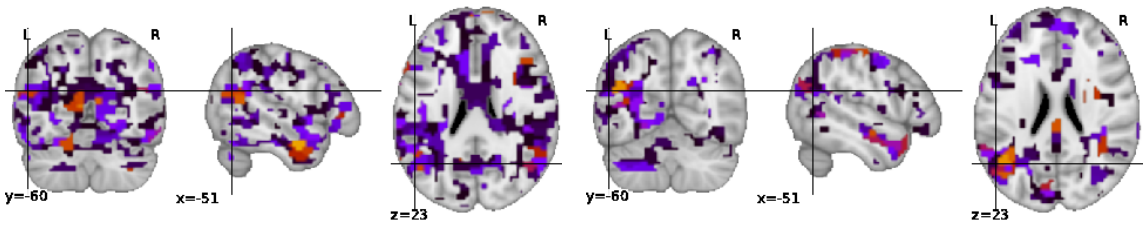


Figure 7. The brain regions that reliably predict that the subject is listening to Korean versus native (french) language (left) are similar to those that can be used to predict that the subject is listening an unintelligible language (jabberwoky) as opposed to their native (french) language (right).

More details can be found in [29] and [30].

6.7. Learning to rank medical images

Participants: Bertrand Thirion, Gaël Varoquaux, Alexandre Gramfort, Fabian Pedregosa [Correspondant].

Medical images can be used to predict a clinical score coding for the severity of a disease, a pain level or the complexity of a cognitive task. In all these cases, the predicted variable has a natural order. While a standard classifier discards this information, we would like to take it into account in order to improve prediction performance. A standard linear regression does model such information, however the linearity assumption is likely not be satisfied when predicting from pixel intensities in an image. In this work we address these modeling challenges with a supervised learning procedure where the model aims to order or rank images. We use a linear model for its robustness in high dimension and its possible interpretation. We show on simulations and two fMRI datasets that this approach is able to predict the correct ordering on pairs of images, yielding higher prediction accuracy than standard regression and multi-class classification techniques.

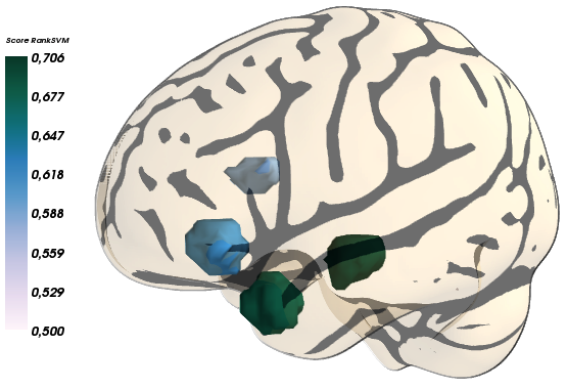


Figure 8. Based on a ranking procedure, the information present in different regions of interest of the brain volume can be used to predict a cognitive feature, in that case the level of complexity of sentences heared by the subject.

More details can be found in [27] and [28].

6.8. Decoding four letter words from brain activations

Participants: Bertrand Thirion, Alexandre Gramfort [Correspondant].

Word reading involves multiple cognitive processes. To infer which word is being visualized, the brain first processes the visual percept, deciphers the letters, bigrams, and activates different words based on context or prior expectation like word frequency. In this contribution, we use supervised machine learning techniques to decode the first step of this processing stream using functional Magnetic Resonance Images (fMRI). We build a decoder that predicts the visual percept formed by four letter words, allowing us to identify words that were not present in the training data. To do so, we cast the learning problem as multiple classification problems after describing words with multiple binary attributes. This work goes beyond the identification or reconstruction of single letters or simple geometrical shapes and addresses a challenging estimation problem, that is the prediction of multiple variables from a single observation, hence facing the problem of learning multiple predictors from correlated inputs.

More details can be found in [22].

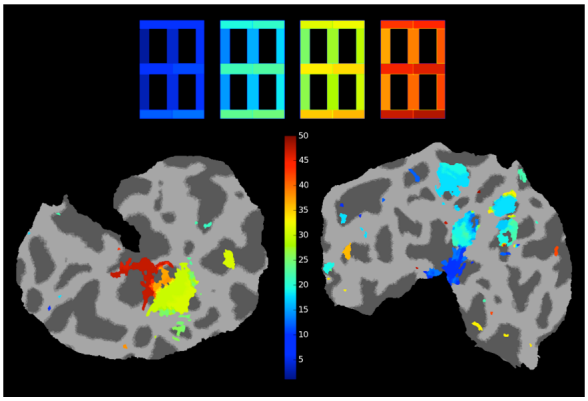


Figure 9. The bars of a word presented on a fixed visual brain activate specific domains of the visual field, and thus can be decoded through this marked. This makes it possible to identify a four letters word presented on a screen.

7. Partnerships and Cooperations

7.1. Regional initiatives

7.1.1. Digiteo/DIM

7.1.1.1. HIDINIM Digiteo project

Participants: Bertrand Thirion [Correspondant], Virgile Fritsch.

High-dimensional Neuroimaging– Statistical Models of Brain Variability observed in Neuroimaging

This is a joint project with Select project team and with SUPELEC Sciences des Systèmes (E3S), Département Signaux & Systèmes Électroniques (A. Tennenhaus).

Statistical inference in a group of subjects is fundamental to draw valid neuroscientific conclusions that generalize to the whole population, based on a finite number of experimental observations. Crucially, this generalization holds under the hypothesis that the population-level distribution of effects is estimated accurately. However, there is growing evidence that standard models, based on Gaussian distributions, do not fit well empirical data in neuroimaging studies.

In particular, Hidinim is motivated by the analysis of new databases hosted and analyzed at Neurospin that contain neuroimaging data from hundreds of subjects, in addition to genetic and behavioral data. We propose to investigate the statistical structure of large populations observed in neuroimaging. In particular, we will investigate the use of region-level averages of brain activity, that we plan to co-analyse with genetic and behavioral information, in order to understand the sources of the observed variability. This entails a series of modeling problems that we will address in this project: i) Distribution normality assessment and variables covariance estimation, ii) model selection for mixture models and iii) setting of classification models for heterogeneous data, in particular for mixed continuous/discrete distributions.

7.1.1.2. ICOGEN Digiteo project

Participants: Bertrand Thirion, Benoit Da Mota [Correspondant].

ICOGEN : Intensive COmputing for GEnetic-Neuroimaging studies

In this project, we design and deploy some computational tools to perform neuroimaging-genetics association studies at a large scale.

Unveiling the relationships between genetic variability and brain structure and function is one of the main challenges in neuroscience, which can be partly addressed through the information conveyed by high-throughput genotyping on the one hand, and neuroimaging data on the other hand. Finding statistical associations between these different variables is important in order to find relevant biomarkers for various brain diseases and improve patient handling. Due to the huge size of the datasets involved and the requirement for tight bounds on statistical significance, such statistical analysis are particularly demanding and cannot be performed easily at a large scale with standard software and computational tools. In ICOGEN, we design and deploy some computational tools to perform neuroimaging-genetics association studies at a large scale. We will implement and assess on real data the use of novel statistical methodologies and run the statistical analysis on various architectures (grids, clouds), in a unified environment.

Project supported by a Digiteo grant in collaboration with Inria's KerData Team, MSR-Inria joint centre, Supélec Engineer School, Imagen project and CEA/Neurospin.

7.1.1.3. SUBSAMPLE Digiteo chair

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Alexandre Abraham.

Parietal is associated with this Digiteo Chair by Dimitris Samaras, in which we will address the probabilistic structure learning of salient brain states (PhD thesis of Alexandre Abraham).

Cognitive tasks systematically involve several brain regions, and exploratory approaches are generally necessary given the lack of knowledge of the complex mechanisms that are observed. The goal of the project is to understand the neurobiological mechanisms that are involved in complex neuro-psychological disorders. A crucial and poorly understood component in this regard refers to the interaction patterns between different regions in the brain. In this project we will develop machine learning methods to capture and study complex functional network characteristics. We hypothesize that these characteristics not only offer insights into brain function but also can be used as concise features that can be used instead of the full dataset for tasks like classification of healthy versus diseased populations or for clustering subjects that might exhibit similarities in brain function. In general, the amount of correlation between distant brain regions may be a more reliable feature than the region-based signals to discriminate between two populations e.g. in Schizophrenia. For such exploratory methods to be successful close interaction with neuroscientists is necessary, as the salience of the features depends on the population and the observed effects of psychopathology. For this aim we propose to develop a number of important methodological advances in the context of prediction of treatment outcomes for drug addicted populations, i.e. for relapse prediction.

7.1.1.4. MMoVNI Digiteo project

Participants: Bertrand Thirion [Correspondant], Pierre Fillard, Viviana Siless, Stéphanie Allassonnière, Hao Xu.

This is a joint project with CMAP http://www.cmapx.polytechnique.fr/~allassonniere/, for the 2010-2013 period.

Modeling and understanding brain structure is a great challenge, given the anatomical and functional complexity of the brain organ. In addition to this, there is a large variability of these characteristics among the population. To give a possible answer to these issues, medical imaging researchers proposed to construct a template image. Most of the time, these analysis only focus on one category of signals (called modality), in particular, the anatomical one was the main focus of research these past years. Moreover, these techniques are often dedicated to a particular problem and raise the question of their mathematical foundations. The MMoVNI project aims at building atlases based on multi-modal images (anatomy, diffusion and functional) data bases for given populations. An atlas is not only a template image but also a set of admissible deformations which characterize the observed population of images. The estimation of these atlases will be based on

a new generation of deformation and template estimation procedures that builds an explicit statistical generative model of the observed data. Moreover, they enable to infer all the relevant variables (parameters of the atlases) thanks to stochastic algorithms. Lastly, this modeling allows also to prove the convergence of both the estimator and the algorithms which provides a theoretical guarantee to the results. The models will first be proposed independently for each modality and then merged together to take into account, in a correlated way, the anatomy, the local connectivity through the cortical fibers and the functional response to a given cognitive task. This model will then be generalized to enable the non-supervised clustering of a population. This leads therefore to a finer representation of the population and a better comparison for classification purposes for example. The Neurospin center, partner of this project, will allow us to have access to databases of images of high-quality and high-resolution for the three modalities: anatomical, diffusion and functional imaging. This project is expected to contribute to making neuroimaging a more reliable tool for understanding inter-subject differences, which will eventually benefit to the understanding and diagnosis of various brain diseases like Alzheimer's disease, autism or schizophrenia.

7.2. National Initiatives

7.2.1. ANR

7.2.1.1. Vimagine project

Participants: Bertrand Thirion [Correspondant], Alexandre Gramfort, Michael Eickenberg, Fabian Pedregosa.

Vimagine is an ANR blanc project (2008-2012), which aims at building a novel view on the retinotopic organization of the visual cortex, based on MEG and MRI. Vimagine should open the way to understanding the dynamics of brain processes for low-level vision, with an emphasis on neuropathologies. This project is leaded by S. Baillet (Dynamic Neuroimaging Lab, McGill University), in collaboration with M.Clerc, T. Papadopoulos (Inria Sophia-Antipolis, Odyssée) and J. Lorenceau(LPPA, CNRS, Collège de France). The fMRI part of the project will be done by PARIETAL, and will consist in a study of spatially resolved retinotopic maps at the mm scale, the decoding of retinotopic information and the comparison of retinotopy with sulcogyral anatomy.

7.2.1.2. BrainPedia project

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Yannick Schwartz, Virgile Fritsch.

BrainPedia is an ANR JCJC (2011-2015) which addresses the following question: Neuroimaging produces huge amounts of complex data that are used to better understand the relations between brain structure and function. While the acquisition and analysis of this data is getting standardized in some aspects, the neuroimaging community is still largely missing appropriate tools to store and organize the knowledge related to the data. Taking advantage of common coordinate systems to represent the results of group studies, coordinate-based meta-analysis approaches associated with repositories of neuroimaging publications provide a crude solution to this problem, that does not yield reliable outputs and looses most of the data-related information. In this project, we propose to tackle the problem in a statistically rigorous framework, thus providing usable information to drive neuroscientific knowledge and questions.

7.2.1.3. IRMgroup project

Participants: Bertrand Thirion [Correspondant], Alexandre Gramfort, Michael Eickenberg.

This is a joint project with Polytechnique/CMAP http://www.cmap.polytechnique.fr/: Stéphanie Allassonnière and Stéphane Mallat (2010-2013).

Much of the visual cortex is organized into visual field maps, which means that nearby neurons have receptive fields at nearby locations in the image. The introduction of functional magnetic resonance imaging (fMRI) has made it possible to identify visual field maps in human cortex, the most important one being the medial occipital cortex (V1,V2,V3). It is also possible to relate directly the activity of simple cells to an fMRI activation pattern and Parietal developed some of the most effective methods. However, the simple cell model is not sufficient to account for high-level information on visual scenes, which requires the introduction of specific semantic features. While the brain regions related to semantic information processing are now well understood, little is known on the flow of visual information processing between the primary visual cortex and the specialized regions in the infero-temporal cortex. A central issue is to better understand the behavior of intermediate cortex layers.

Our proposition is to use our mathematical approach to formulate explicitly some generative model of information processing, such as those that characterize complex cells in the visual cortex, and then to identify the brain substrate of the corresponding processing units from fMRI data. While fMRI resolution is still too coarse for a very detailed mapping of detailed cortical functional organization, we conjecture that some of the functional mechanisms that characterize biological vision processes can be captured through fMRI; in parallel we will push the fMRI resolution to increase our chance to obtain a detailed mapping of visual cortical regions.

7.2.1.4. Niconnect project

Participants: Bertrand Thirion, Gaël Varoquaux [Correspondant], Alexandre Abraham.

- **Context:** The project NiConnect arises from an increasing need of medical imaging tools to diagnose efficiently brain pathologies, such as neuro-degenerative and psychiatric diseases or lesions related to stroke. Brain imaging provides a non-invasive and widespread probe of various features of brain organization, that are then used to make an accurate diagnosis, assess brain rehabilitation, or make a prognostic on the chance of recovery of a patient. Among different measures extracted from brain imaging, functional connectivity is particularly attractive, as it readily probes the integrity of brain networks, considered as providing the most complete view on brain functional organization.
- **Challenges:** To turn methods research into popular tool widely usable by non specialists, the NiConnect project puts specific emphasis on producing high-quality open-source software. NiConnect addresses the many data analysis tasks that extract relevant information from resting-state fMRI datasets. Specifically, the scientific difficulties are i) conducting proper validation of the models and tools, and ii)providing statistically controlled information to neuroscientists or medical doctors. More importantly, these procedures should be robust enough to perform analysis on limited quality data, as acquiring data on diseased populations is challenging and artifacts can hardly be controlled in clinical settings.
- **Outcome of the project:** In the scope of computer science and statistics, NiConnect will push forward algorithms and statistical models for brain functional connectivity. In particular, we are investigating structured and multi-task graphical models to learn high-dimensional multi-subject brain connectivity models, as well as spatially-informed sparse decompositions for segmenting structured from brain imaging. With regards to neuroimaging methods development, NiConnect will provide systematic comparisons and evaluations of connectivity biomarkers and a software library embedding best-performing state-of-the-art approaches. Finally, with regards to medical applications, the NiConnect project will also play a support role in on going medical studies and clinical trials on neurodegenerative diseases.
- **Consortium**
  - Parietal Inria research team: applied mathematics and computer science to model the brain from MRI
  - LIF INSERM research team: medical image data analysis and modeling for clinical applications
  - CATI center: medical image processing center for large scale brain imaging studies

- Henri-Mondor hospital neurosurgery and neuroradiology: clinical teams conducting research on treatments for neurodegenerative diseases, in particular Huntington and Parkinson diseases
- Logilab: consulting in scientific computing

7.3. International Initiatives

7.3.1. Inria Associate Teams

Title: Analysis of structural MR and DTI in neonates

Inria principal investigator: Pierre Fillard

International Partner:

Institution: University of Southern California (United States)

Laboratory: Image Lab at Children Hospital at Los Angeles

Researcher: Natasha Lepore

International Partner:

Institution: University of Pennsylvania (United States)

Laboratory: Penn Image Computing and Science Laboratory

Researcher: Caroline Brun

Duration: 2011 - 2013

See also: http://www.capneonates.org/

While survival is possible at increasingly lower gestational ages at birth, premature babies are at higher risk of developing mental disorders or learning disabilities than babies born at term. A precise identification of the developmental differences between premature and control neonates is consequently of utmost importance. Nowadays, the continuously improving quality and availability of MR systems makes it possible to precisely determine, characterize and compare brain structures such as cortical regions, or white matter fiber bundles. The objective of this project is to understand the developmental differences of premature versus normal neonates, using structural and diffusion MRI. This work will consist in identifying, characterizing and meticulously studying the brain structures that are different between the two groups. To do so, we propose to join forces between the Parietal team at Inria and the University of Southern California. Parietal has a recognized expertise in medical image registration and in statistical analyses of groups of individuals. USC has a broad knowledge in MR image processing. In particular, the Children's Hospital at Los Angeles (CHLA), which is part of USC, is in the process of collecting a unique database of several hundreds of premature and normal neonates MR scans. This joint collaboration is consequently a unique chance of addressing key questions pertaining to neonatal and premature development. It will make it possible to elaborate new tools to analyze neonate MR images while tremendously increasing our knowledge of neuroanatomy at such an early stage in life.

7.3.2. Inria International Partners

- LIAMA http://www.nlpr.ia.ac.cn/jiangtz/: B.Thirion, G.Varoquaux, V. Siless and Y. Schwartz visited LIAMA (contact person: Shan Yu) and gave a presentation. We plan to develop come collaborations on fMRI data analysis and functional connectivity in the future.
- Donders institute https://sites.google.com/a/distrep.org/distrep/marcel-van-gerven: We share with M. van Gerven some interest on biological vision and on the use of fMRI to probe specific hypotheses related to computational models of vision. We plan to exchange students in the next years.
- Biomedical Image analysis group, Imperial College, London http://www.doc.ic.ac.uk/~dr/. We have started some joint work on the comparison of functional and anatomical connectivity using machine learning tools.

- MIT, CSAIL http://www.csail.mit.edu/, P.Golland's group. We regularly visit each other and share common interests in the use of machine learning for neuroimaging, in the introduction of functional information into co-registration procedures, and in the study and comparison of anatomical and functional connectivity.

7.3.3. Participation In International Programs

Parietal has taken part to the program Inria@SiliconValley, and had a 18-months post-doc funded to work on the comparison of anatomical and functional connectivity (18 months, 2011-2013):

In this project, we would like to build probabilistic models that relates quantitatively the observations in anatomical and functional connectivity. For instance given a set of brain regions, the level of functional integration might be predicted by the anatomical connectivity measurement derived from the fibers in a given population of subjects. More generally, we will seek to extract latent factors explaining both connectivity measures across the population. Such models require specifically that a generative model is proposed to explain the observations in either domain, so that a meaningful and testable link is built between the two modalities. The inference problem can then be formulated as learning the coupling parameters that are necessary to model the association between modalities, and tested e.g. by assessing the ability of the learned model to generalize to new subjects. The aim is then to provide the mathematical and algorithmic tools necessary to build a standardized model of brain connectivity informed by both modalities, associated with confidence intervals to take into account between subject variability. Such an atlas is a long-term project, that requires adequate validation on high-resolution data, but it will probably be tightly linked to this project.

8. Dissemination

8.1. Scientific Animation

- B. Thirion acts as reviewers for Medical Image Analysis, IEEE Transactions on Medical Imaging, NeuroImage, ISBI, IPMI, as associate editor for Frontiers in Neuroscience Methods, as program committee for the MICCAI 2012 conference and as expert for ANR, NWO.
- B.Thirion set up the following workshop at the OHBM 2012 conference: *Why believe in Multivariate Pattern Analysis ? The skeptical Neuroimager's view.* G. Varoquaux was part of it. B.Thirion organized a session entitled *Comprendre le cerveau par le signal et l'image* at the Digiteo forum.
- B. Thirion gave an invited presentation at the ISBI 2012, CompImage 2012, the FENS 2012 workshop on brain activity decoding, MICCAI 2012 workshop on functional connectivity and at the MLINI 2012 workshop, at the *Institut de biologie du Collège de France*, Paris and at the *Institute of Psychiatry*, London.
- G. Varoquaux gave a tutorial on the use of Python in Machine learning at the PRNI 2012 conference.
- G. Varoquaux was program committee of Euroscipy 2012, PyHPC 2012 and ESCO 2012.
- G. Varoquaux acts as reviewer for NeuroImage, HBM, MedIA, TMI, Frontiers in NeuroInformatics et Frontiers in Brain Imaging methods, Review editor for Frontiers in NeuroInformatics and Frontiers in Brain Imaging methods and as expert for ANR and Agoranov.
- Alexandre Gramfort is Program committee PRNI, Associate editor IEEE EMBC conference and Associate editor Frontiers in brain imaging methods.
- Alexandre Gramfort acts as reviewer for Neuroimage, IEEE TMI, brain topography, HBM journal, PLOS ONE, brain connectivity, journal of clinical neurophysiology, MICCAI, physics in medicine and biology
- Alexandre Gramfort took part to the *Inria-Clinatec* days at Grenoble.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master: MVA, Bertrand Thirion + Alexandre Gramfort, Imagerie fonctionnelle cérébrale et interface cerveau machine, 12h + 3h, M2, ENS Cachan, France.

Master biostatistiques: Bertrand Thirion, cours de biostatistique computationnelle, 6h, M2, Paris XI, France.

PhD: Gaël Varoquaux taught Python (8h at Inria Saclay, 8h at Inria Grenoble, 4h at Expresso summer school).

Master: master Telecom ParisTech UE ACIMED, Alexandre Gramfort tought MEG + fMRI 4h30, M2, Telecom ParisTech, Paris, France

8.2.2. Juries

Bertrand Thirion was part of the PhD committee of

- Peter Rasmussen, DTU, Denmark on April 11th. The PhD thesis was entitled *Mathematical modeling and visualization of functional neuroimages*.
- Archana Venkataraman, MIOT/CSAIL, USA on June 26th. The PhD thesis was entitled *Generative Models of Brain Connectivity for Population Studies*.
- Mikael Naveau, Cyceron, Caen, France on November 5th. The PhD thesis was entitled *Connectivité fonctionnelle cérébrale pendant l'état de repos: modélisation multi-échelle*.

8.3. Popularization

Alexandre Gramfort was shot in the *E=M6* TV program, in which Marc lesguy took part to the so-called word decoding experiment.

Parietal animated the Inria stand at Fête de la science days, Moulon Saclay, during 3 days (October 12-14), where we presented a nice game designed by Virgile Fritsch to illustrate our research activities.

9. Bibliography

Major publications by the team in recent years

[1] P. FILLARD, C. POUPON, J.-F. MANGIN. *A Novel Global Tractography Framework based on an Adaptive Spin Glass Model*, in "Proc. 12th MICCAI", 2009, ftp://static.lnao.fr/lnao/static/papers/Fillard-MICCAI09.pdf.

[2] V. FRITSCH, G. VAROQUAUX, B. THYREAU, J.-B. POLINE, B. THIRION. *Detecting Outliers in High-Dimensional Neuroimaging Datasets with Robust Covariance Estimators*, in "Medical Image Analysis", May 2012, vol. 16, p. 1359-1370 [*DOI :* 10.1016/J.MEDIA.2012.05.002], http://hal.inria.fr/hal-00701225.

[3] R. JENATTON, A. GRAMFORT, V. MICHEL, G. OBOZINSKI, E. EGER, F. BACH, B. THIRION. *Multi-scale Mining of fMRI data with Hierarchical Structured Sparsity*, in "SIAM Journal on Imaging Sciences", July 2012, vol. 5, n^{o} 3, p. 835-856 [*DOI :* 10.1137/110832380], http://hal.inria.fr/inria-00589785.

[4] A. KNOPS, B. THIRION, E. HUBBARD, V. MICHEL, S. DEHAENE. *Recruitment of an area involved in eye movements during mental arithmetic.*, in "Science", Jun 2009, vol. 324, n^{o} 5934, p. 1583–1585.

[5] V. MICHEL, A. GRAMFORT, G. VAROQUAUX, E. EGER, B. THIRION. *Total variation regularization for fMRI-based prediction of behaviour.*, in "IEEE Transactions on Medical Imaging", February 2011, vol. 30, n^{o} 7, p. 1328 - 1340 [*DOI :* 10.1109/TMI.2011.2113378], http://hal.inria.fr/inria-00563468/en.

[6] G. VAROQUAUX, F. BARONNET, A. KLEINSCHMIDT, P. FILLARD, B. THIRION. *Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling*, in "Medical Image Computing and Computer Added Intervention", Chine Beijing, Springer, Sep 2010, http://hal.inria.fr/inria-00512417.

[7] G. VAROQUAUX, A. GRAMFORT, F. PEDREGOSA, V. MICHEL, B. THIRION. *Multi-subject dictionary learning to segment an atlas of brain spontaneous activity*, in "Information Processing in Medical Imaging", Kaufbeuren, Germany, Lecture Notes in Computer Science, Springer, July 2011, vol. 6801, p. 562-573 [*DOI :* 10.1007/978-3-642-22092-0_46], http://hal.inria.fr/inria-00588898/en.

[8] G. VAROQUAUX, A. GRAMFORT, J.-B. POLINE, B. THIRION. *Brain covariance selection: better individual functional connectivity models using population prior*, in "Advances in Neural Information Processing Systems", Canada Vancouver, John Lafferty, Dec 2010, http://hal.inria.fr/inria-00512451.

[9] G. VAROQUAUX, A. GRAMFORT, B. THIRION. *Small-sample brain mapping: sparse recovery on spatially correlated designs with randomization and clustering*, in "International Conference on Machine Learning", Edimbourg, United Kingdom, L. JOHN, P. JOELLE (editors), Andrew McCallum, June 2012, http://hal.inria.fr/hal-00705192.

Publications of the year

Articles in International Peer-Reviewed Journals

[10] G. ANTONIU, A. COSTAN, B. DA MOTA, B. THIRION, R. TUDORAN. *A-Brain: Using the Cloud to Understand the Impact of Genetic Variability on the Brain*, in "ERCIM News", April 2012, p. 21-22, http://hal.inria.fr/hal-00684384.

[11] V. FRITSCH, G. VAROQUAUX, B. THYREAU, J.-B. POLINE, B. THIRION. *Detecting Outliers in High-Dimensional Neuroimaging Datasets with Robust Covariance Estimators*, in "Medical Image Analysis", May 2012, vol. 16, p. 1359-1370 [*DOI :* 10.1016/J.MEDIA.2012.05.002], http://hal.inria.fr/hal-00701225.

[12] E. GOUILLART, M.-J. TOPLIS, J. GRYNBERG, M.-H. CHOPINET, E. SONDERGARD, L. SALVO, M. SUÉRY, M. DI MICHIEL, G. VAROQUAUX. *In Situ Synchrotron Microtomography Reveals Multiple Reaction Pathways During Soda-Lime Glass Synthesis*, in "Journal of the American Ceramic Society", 2012, vol. 95, n^{o} 5, p. 1504-1507 [*DOI :* 10.1111/J.1551-2916.2012.05151.X], http://hal.inria.fr/hal-00705955.

[13] A. GRAMFORT, M. KOWALSKI, M. HÄMÄLÄINEN. *Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods*, in "Physics in Medicine and Biology", March 2012, vol. 57, n^{o} 7, p. 1937-1961 [*DOI :* 10.1088/0031-9155/57/7/1937], http://hal.inria.fr/hal-00690774.

[14] P. GUEVARA, D. DUCLAP, C. POUPON, L. MARRAKCHI-KACEM, P. FILLARD, D. LE BIHAN, M. LEBOYER, J. HOUENOU, J.-F. MANGIN. *Automatic fiber bundle segmentation in massive tractography datasets using a multi-subject bundle atlas.*, in "NeuroImage", March 2012, p. 1083-99 [*DOI :* 10.1016/J.NEUROIMAGE.2012.02.071], http://hal.inria.fr/hal-00700800.

[15] R. JENATTON, A. GRAMFORT, V. MICHEL, G. OBOZINSKI, E. EGER, F. BACH, B. THIRION. *Multi-scale Mining of fMRI data with Hierarchical Structured Sparsity*, in "SIAM Journal on Imaging Sciences", July 2012, vol. 5, n^{o} 3, p. 835-856 [*DOI :* 10.1137/110832380], http://hal.inria.fr/inria-00589785.

[16] A. TUCHOLKA, V. FRITSCH, J.-B. POLINE, B. THIRION. *An empirical comparison of surface-based and volume-based group studies in neuroimaging*, in "NeuroImage", June 2012, p. 1443-53 [*DOI :* 10.1016/J.NEUROIMAGE.2012.06.019], http://hal.inria.fr/hal-00723437.

[17] G. VAROQUAUX, A. GRAMFORT, J.-B. POLINE, B. THIRION. *Markov models for fMRI correlation structure: is brain functional connectivity small world, or decomposable into networks?*, in "Journal of Physiology - Paris", January 2012, p. 212-221 [*DOI :* 10.1016/J.JPHYSPARIS.2012.01.001], http://hal.inria.fr/hal-00665340.

International Conferences with Proceedings

[18] B. DA MOTA, M. EICKENBERG, S. LAGUITTON, V. FROUIN, G. VAROQUAUX, J.-B. POLINE, B. THIRION. *A MapReduce Approach for Ridge Regression in Neuroimaging-Genetic Studies*, in "DCICTIA-MICCAI - Data- and Compute-Intensive Clinical and Translational Imaging Applications in conjonction with the 15th International Conference on Medical Image Computing and Computer Assisted Intervention - 2012", Nice, France, October 2012, http://hal.inria.fr/hal-00730385.

[19] B. DA MOTA, V. FROUIN, E. DUCHESNAY, S. LAGUITTON, G. VAROQUAUX, J.-B. POLINE, B. THIRION. *A fast computational framework for genome-wide association studies with neuroimaging data*, in "20th International Conference on Computational Statistics (COMPSTAT 2012)", Limassol, Cyprus, 2012, http://hal.inria.fr/hal-00720265.

[20] M. EICKENBERG, A. GRAMFORT, B. THIRION. *Multilayer Scattering Image Analysis Fits fMRI Activity in Visual Areas*, in "International Workshop on Pattern Recognition in NeuroImaging", London, United Kingdom, July 2012, http://hal.inria.fr/hal-00704528.

[21] V. FRITSCH, G. VAROQUAUX, J.-B. POLINE, B. THIRION. *Non-parametric Density Modeling and Outlier Detection in Medical Imaging Datasets*, in "Machine Learning in Medical Imaging - Miccai 2012 workshop", Nice, France, F. WANG, D. SHEN, P. YAN, K. SUZUKI (editors), October 2012, p. 207-214, http://hal.inria.fr/hal-00738438.

[22] A. GRAMFORT, C. PALLIER, G. VAROQUAUX, B. THIRION. *Decoding Visual Percepts Induced by Word Reading with fMRI*, in "Pattern Recognition in NeuroImaging (PRNI), 2012 International Workshop on", Londres, United Kingdom, July 2012, p. 13-16 [*DOI :* 10.1109/PRNI.2012.20], http://hal.inria.fr/hal-00730768.

[23] A. GRAMFORT, C. POUPON, M. DESCOTEAUX. *Sparse DSI: Learning DSI structure for denoising and fast imaging*, in "MICCAI", Nice, France, Springer, October 2012, http://hal.inria.fr/hal-00723897.

[24] N. LEPORE, F. YEPES, Y. LAO, A. PANIGRAPHY, R. CHESCHIN, S. RAVICHANDRAN, M. D. NELSON, P. FILLARD. *Template-Based Tractography for Clinical Neonatal Diffusion Imaging Data*, in "Template-Based Tractography for Clinical Neonatal Diffusion Imaging Data", San Francisco CA, United States, SPIE, January 2012, http://hal.inria.fr/hal-00675337.

[25] B. NG, V. SILESS, G. VAROQUAUX, J.-B. POLINE, B. THIRION, R. ABUGHARBIEH. *Connectivity-informed Sparse Classifiers for fMRI Brain Decoding*, in "Pattern Recognition in Neuroimaging", London, United Kingdom, Christos Davatzikos, Moritz Grosse-Wentrup, Janaina Mourao-Miranda, Dimitri Van De Ville, July 2012, http://hal.inria.fr/hal-00726656.

[26] B. NG, G. VAROQUAUX, J.-B. POLINE, B. THIRION. *A Novel Sparse Graphical Approach for Multimodal Brain Connectivity Inference*, in "Medical Image Computing and Computer Assisted Intervention", Nice, France, October 2012, http://hal.inria.fr/hal-00741631.

[27] F. PEDREGOSA, A. GRAMFORT, G. VAROQUAUX, E. CAUVET, C. PALLIER, B. THIRION. *Learning to rank from medical imaging data*, in "MLMI 2012 - 3rd International Workshop on Machine Learning in Medical Imaging", Nice, France, Inria, July 2012, http://hal.inria.fr/hal-00717990.

[28] F. PEDREGOSA, A. GRAMFORT, G. VAROQUAUX, B. THIRION, C. PALLIER, E. CAUVET. *Improved brain pattern recovery through ranking approaches*, in "PRNI 2012 : 2nd International Workshop on Pattern Recognition in NeuroImaging", London, United Kingdom, July 2012, http://hal.inria.fr/hal-00717954.

[29] Y. SCHWARTZ, G. VAROQUAUX, C. PALLIER, P. PINEL, J.-B. POLINE, B. THIRION. *Improving accuracy and power with transfer learning using a meta-analytic database*, in "MICCAI", Nice, France, October 2012, p. 1-8, http://hal.inria.fr/hal-00734911.

[30] Y. SCHWARTZ, G. VAROQUAUX, B. THIRION. *On spatial selectivity and prediction across conditions with fMRI*, in "PRNI 2012 : 2nd International Workshop on Pattern Recognition in NeuroImaging", London, United Kingdom, July 2012, p. 53-56, http://hal.inria.fr/hal-00728765.

[31] V. SILESS, J. GLAUNÈS, P. GUEVARA, J.-F. MANGIN, C. POUPON, D. LE BIHAN, B. THIRION, P. FILLARD. *Joint T1 and Brain Fiber Log-Demons Registration Using Currents to Model Geometry*, in "MICCAI", Nice, France, France, October 2012, http://hal.inria.fr/hal-00723367.

[32] G. VAROQUAUX, A. GRAMFORT, B. THIRION. *Small-sample brain mapping: sparse recovery on spatially correlated designs with randomization and clustering*, in "International Conference on Machine Learning", Edimbourg, United Kingdom, L. JOHN, P. JOELLE (editors), Andrew McCallum, June 2012, http://hal.inria.fr/hal-00705192.

References in notes

[33] K. JIMURA, R. A. POLDRACK. *Analyses of regional-average activation and multivoxel pattern information tell complementary stories*, in "Neuropsychologia", 2012, vol. 50, 544.